

Temporal code versus rate code for binary Information Sources

Agnieszka Pregowska^a, Janusz Szczepanski^{a,*}, Eligiusz Wajnryb^a

^a*Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawlowskiego
5B, Warsaw*

Abstract

Neuroscientists formulate very different hypotheses about the nature of neural code. At one extreme, it has been argued that neurons encode information in relatively slow changes of individual spikes arriving rates (rates codes) and the irregularity in the spike trains reflects noise in the system, while in the other extreme this irregularity is the code itself (temporal codes) thus the precise timing of every spike carries additional information about the input. It is well known that in the estimation of Shannon information transmission rate the patterns and temporal structures are taken into account, while the "rate code" is already determined by the firing rate, i.e. by spike frequency. In this paper we compare these two types of codes for binary Information Sources which model encoded spike-trains. Assuming that the information transmitted by a neuron is governed by uncorrelated stochastic process or by process with a memory we compare the information transmission rates carried by such spike-trains with their firing rates. We showed that the crucial role in studying the relation between information and firing rates is played by a quantity which we call "jumping" parameter. It corresponds to the probabilities of transitions from no-spike-state to the spike-state and vice versa. For low values of jumping parameter the quotient of information and firing rates is monotonically decreasing function of firing rate, thus there is straightforward, one-to-one, relation between temporal

*Corresponding author

Email address: jszczepa@ippt.pan.pl (Janusz Szczepanski)

and rate codes. On the contrary, it turns out that for large enough jumping parameter this quotient is non-monotonic function of firing rate and it exhibits a global maximum, i.e. in this case there exists the optimal firing rate. Moreover, there is no one-to-one relation between information and firing rates, so the temporal and rate codes differ qualitatively. This leads to the observation that the behavior of the quotient of information and firing rates for large jumping parameter is especially important in the context of bursting phenomena.

Keywords: Information Theory, Information Source, stochastic process, information transmission rate, firing rate

1. Introduction

Fundamental Neuroscience problem is to understand how neurons encode and process information [21, 23, 24]. In general it is not easy to determine the neural code structure. Since Adrian's experiments [1] which established that individual sensory neurons produce action potentials, or spikes it is assumed that a single neuron provides information just through spikes sequence, i.e. spike-trains. Although it is now generally accepted that a spike sequence is the way the information is coded by a single neuron, the structure and the mechanisms of code formation are still mysteries. In 1976 Burns and Webb [6] for the first time showed that the total number of emitted spikes arrives in a highly irregular manner. When the same stimulus is applied repeatedly the number of spikes varies substantially from trial to trial [20]. This has inclined Neuroscientists to formulate very different hypotheses about the nature of the neural code. Two main ideas, not excluding each other, are of special interest. The first theory is based on the idea of "temporal code" [8, 10, 24, 25] and goes into the spike -trains structure while the second referred to as "rate code" theory [9, 15, 19, 23, 24] assumes that the neural code is embedded in the spike frequency, defined as the number of spikes emitted per second. The temporal coding mechanism, which builds a relationship of temporal process between the output firing patterns and the inputs of the nervous system, has received

much attention [7, 11, 13]. On the other hand in the transfer of information the most expansive energetically is spiking process [2, 16], thus in the first approximation the firing rate can be treated as the energy marker. Inspired by the thermodynamics [33] we also consider the derivative of entropy over energy which is the analog of temperature inverse.

In this article we give the theoretical insight into understanding the neural code nature, namely we study this problem for two types of binary Information Sources. Assuming that the information transmitted by a neuron is governed by uncorrelated stochastic process or by process with a memory we study the relation between the Information Transmission Rates ITR carried by such spike-trains and their firing rate F_R . To this end the Information-Firing-Quotient IFQ , being the ratio of information and firing rate, is introduced in Section 2. For large IFQ transmission is more optimal in the sense of information amount transmitted at the cost of unit energy. We show that the crucial role in studying IFQ properties is played by the "jumping" parameter. This parameter is the sum of transition probabilities from no-spike-state to spike-state and vice versa. We show that for the low values of jumping parameter the quotient of information and firing rates is monotonically decreasing function of firing rate, thus there is straightforward, one-to-one, relation between temporal and rate codes. On the contrary, it turns out that for large enough jumping parameter this quotient is non-monotonic function of firing rate and it exhibits well pronounced global maximum. Thus, in this case the optimal firing rate exists. Moreover, there is no one-to-one relation between information and firing rate and the temporal and rate codes differ qualitatively. The behavior of the quotient of information and firing rates for large jumping parameter is especially important in the context of bursting phenomenon[34–36].

The paper is organized as follows. In Section 2 the basic concepts of Information Theory and formulas concerning Bernoulli and Markov processes are briefly recalled. The comparison of information transmission and firing rates for these processes is presented in Section 3. The last Section contains discussion and conclusions.

2. Information Theory in Neuroscience

In Neuroscience the information transfer is quantified by many authors in terms of Information Theory [5, 24]. In general, neuronal communication systems are represented by Information Source, communication channel and output signals [22, 26, 27]. Both messages coming from Information Source and output signals are represented by sequences of symbols [4, 5, 24, 29, 34]. These sequences can be understood as trajectories of stationary stochastic processes. In this paper we study the Information Sources which are represented by Bernoulli or Markov processes [4, 12].

Entropy. First, we briefly recall the fundamental concepts of Information Theory [22, 26, 27] that are adapted to Neuroscience issues. Let Z^L be a set of all words (i.e. blocks) of length L , built of symbols (letters) from some finite alphabet Z . Each word z^L can be treated as a message sent by Information Source Z being a stationary stochastic process. If $P(z^L)$ denotes the probability the word $z^L \in Z^L$ occurs, then the information in Shannon sense carried by this word is defined as

$$I(z^L) := -\ln P(z^L) . \quad (1)$$

In this sense, less probable events carry more information. We use the natural logarithm to get more compact form of the formulas. In case when logarithm to the base 2 is used, the factor of $\ln 2$ appears. Expected or average information of Z^L , called Shannon block entropy reads

$$H(Z^L) := - \sum_{z^L \in Z^L} P(z^L) \ln P(z^L) . \quad (2)$$

Since the word length L can be chosen arbitrary, the block entropy does not perfectly describe the Information Source [26, 27]. The more adequate characteristics of the information transmission rate is defined in the next Subsection. For the special case of two-letter alphabet $Z = \{0, 1\}$ and the length of words $L = 1$ we introduce the following notation for the entropy

$$H_1(p) := H(Z^1) = -p \ln p - (1 - p) \ln (1 - p) , \quad (3)$$

where $P(1) = p, P(0) = 1 - p$ are the associated probabilities. This is, in fact, formula for the entropy of two-state system.

Information Transmission Rate. The entropy of spike trains themselves evaluates how much information these spikes could provide. The adequate measure for estimation of efficiency of information source is the information transmitted in average by a single symbol. This measure, which characterizes Information Source $\{Z\}$, is called Information Transmission Rate ITR and is defined as [26, 27]

$$ITR(\{Z\}) := \lim_{L \rightarrow \infty} \frac{H(Z^L)}{L} . \quad (4)$$

Information transmission rate exists if and only if the stochastic process $\{Z\}$ is stationary [27].

The Information Transmission Rate is very important quantity especially due to the Asymptotic Equipartition Theorem. This theorem states that information per symbol for most of the messages coming from a given source is close to ITR [26, 27].

Firing Rate. Since the experiment of Adrian [1] very important characterization of both neural network dynamics and neural computation is the firing rate F_R of spike-trains. The first and most commonly used definition of the firing rate refers to temporal average [11, 24, 34] and reads

$$F_R = \frac{n_T}{T} , \quad (5)$$

where n_T denotes spike count and T is time window length. In practice, in order to get sensible averages, some reasonable number of spikes should occur within the time window. Since the messages are trajectories of stationary stochastic process the firing rate as defined by (5) is specific for a given Information Source provided T is large enough. Thus, $F_R \cdot \Delta\tau$ can be identified with the probability p of spike appearance, where $\Delta\tau$ is the time resolution or bin size.

Bernoulli process. Assuming the size of bin spike trains can be encoded [4] in such a way that 1 is generated with probability p (spike is arrived in the bin),

0 is generated with probability $1 - p$ (spike is not arrived). In the situation when consecutive bits in message are uncorrelated, we are in the regime of Bernoulli process [12, 27]. Following the entropy definition (2) the Information Transmission Rate (5) for Bernoulli process reads

$$ITR(p) = -p \ln p - (1 - p) \ln (1 - p) = H_1(p) . \quad (6)$$

Further in the paper the Bernoulli process will be considered as a benchmark for more complex processes like the Markov ones.

Markov process. Again assuming the size of bin we now consider as the Information Source the discrete-time, two-state Markov processes. The conditional probabilities for such processes are completely defined in terms of two transition probabilities from state 0 to state 1, $p_{1|0}$ and from state 1 to state 0, $p_{0|1}$. The Markov transition probability matrix P can be written as

$$P := \begin{bmatrix} 1 - p_{1|0} & p_{0|1} \\ p_{1|0} & 1 - p_{0|1} \end{bmatrix} . \quad (7)$$

We assume here that the process is homogeneous in time.

The probability evolution is governed by Master Equation [33]

$$\begin{bmatrix} P_{n+1}(0) \\ P_{n+1}(1) \end{bmatrix} = \begin{bmatrix} 1 - p_{1|0} & p_{0|1} \\ p_{1|0} & 1 - p_{0|1} \end{bmatrix} \cdot \begin{bmatrix} P_n(0) \\ P_n(1) \end{bmatrix} , \quad (8)$$

where n stands for the discrete time, and the stationary solution reads

$$\begin{bmatrix} P_{eq}(0) \\ P_{eq}(1) \end{bmatrix} = \begin{bmatrix} p_{0|1}/(p_{0|1} + p_{1|0}) \\ p_{1|0}/(p_{0|1} + p_{1|0}) \end{bmatrix} . \quad (9)$$

The Information Transmission Rate (4) of such a Markov source reads [27].

$$\begin{aligned} ITR = & P_{eq}(0)(-p_{1|0} \ln p_{1|0} - (1 - p_{1|0}) \ln (1 - p_{1|0})) \\ & + P_{eq}(1)(-p_{0|1} \ln p_{0|1} - (1 - p_{0|1}) \ln (1 - p_{0|1})) , \end{aligned} \quad (10)$$

or making use of notation (3), in compact form

$$ITR = P_{eq}(0)H_1(p_{1|0}) + P_{eq}(1)H_1(p_{0|1}) . \quad (11)$$

For the later use the probability of state "1" is for short denoted by p

$$p := P_{eq}(1) = \frac{p_{1|0}}{(p_{0|1} + p_{1|0})} . \quad (12)$$

and in fact is interpreted as the firing rate. Please, note that for the special case when $p_{0|1} + p_{1|0} = 1$ the Markov process becomes uncorrelated and reduces to Bernoulli process with $p = p_{1|0}$.

3. Results

It is well known that the "temporal code" approach requires the reliable estimation of Information Transmission Rate which must take into account the patterns and temporal structures [18, 28? –32], while the "rate code" is determined just by the firing rate, which in turn is fully given by the probability p . Addressing the problem of relation between "temporal code" and "rate code" we introduce the Information-Firing-Quotient IFQ defined as the ratio

$$IFQ := \frac{ITR}{F_R} . \quad (13)$$

IFQ can be understood as the information cost in terms of the energy units. Further we analyze the IFQ for messages coming from two qualitatively different Information Sources, namely Bernoulli and Markov processes.

3.1. Information Source of Bernoulli Type

Let us consider the Information-Firing-Quotient formula for Information Source being a Bernoulli process with probability parameter p . For this source we denote the IFQ by $B(p)$.

Using (4) and (6) we get

$$B(p) := \frac{H_1(p)}{p} \quad (14)$$

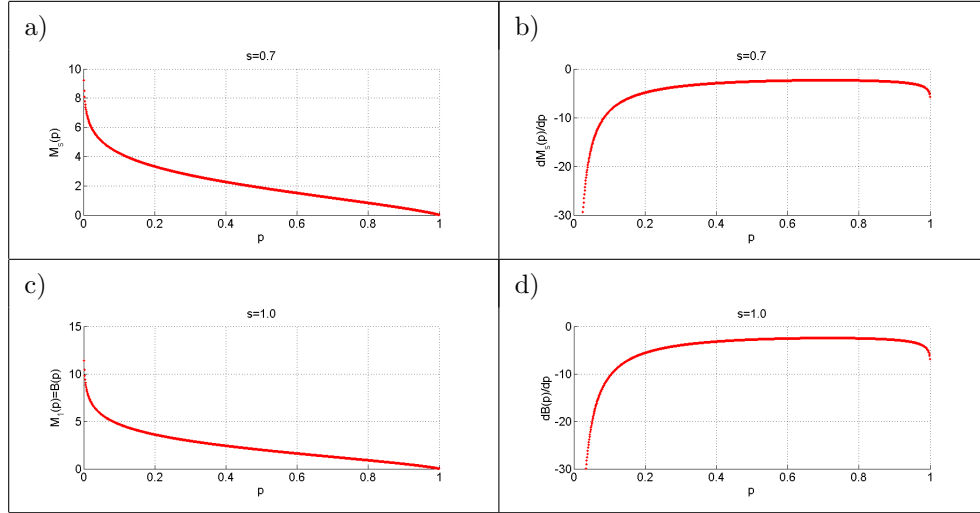


Figure 1: The typical course of Information-Firing-Quotient M_s versus spiking probability p for low values of jumping parameter $s \leq 1$. In panels a, b course of M_s and its derivative for $s = 0.7$ is shown. Panels c, d the course of M_s and its derivative for $s = 1.0$ is presented. Notice that $s = 1$ is the Bernoulli source case.

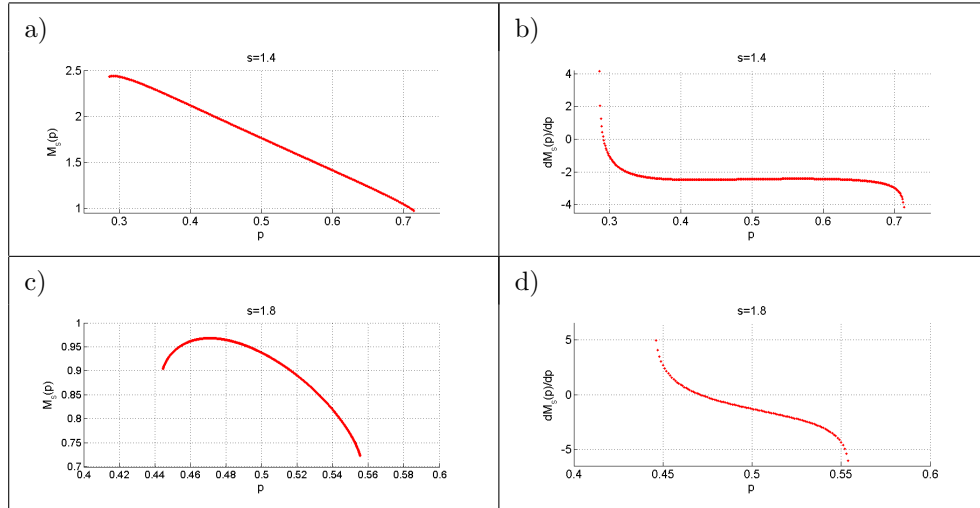


Figure 2: The typical course of Information-Firing-Quotient M_s versus spiking probability p for larger jumping parameter $s > 1$. Panels a, b course of M_s and its derivative for $s = 1.4$ is shown. Panels c, d the course of M_s and its derivative for more jumping case i.e. $s = 1.8$ is shown. Observe, that the graphs for $s > 1$ are qualitatively different than for $s < 1$, see Fig. 1.

for $0 < p \leq 1$. Now, we evaluate the derivative of the quotient $B(p)$ over the firing rate p (corresponding to energy), which has the form

$$\frac{dB}{dp}(p) = \frac{d}{dp} \left(\frac{ITR(p)}{f_R(p)} \right) = \frac{1}{p^2} \ln(1-p) \quad (15)$$

for $0 < p \leq 1$. In order to find lower and upper bounds of these expressions we make use of the following inequalities

$$1 - \frac{1}{x} \leq \ln x \leq x - 1, \quad (16)$$

which hold for $0 \leq x$ and the inequality

$$4(1-x)x \ln 2 \leq H_1(x) \leq \ln 2, \quad (17)$$

which is true for $0 < x \leq 1$.

Applying (16) and (17) to (14) we obtain the following bounds for $B(p)$

$$4(1-p) \leq B(p) \ln 2 \leq \frac{\ln 2}{p} \quad (18)$$

and making use of (15) and (16) we get bound of the derivative of $B(p)$

$$-\frac{1}{(1-p)p} \leq \frac{dB}{dp}(p) \leq -\frac{1}{p}. \quad (19)$$

Introducing the following notation

$$L_0(p) := 4(1-p), \quad (20)$$

$$U_0(p) := \frac{1}{p}, \quad (21)$$

$$L_1(p) := \frac{1}{(1-p)p}, \quad (22)$$

we express (17) and (18) in the compact forms

$$L_0(p) \ln 2 \leq B(p) \leq U_0(p) \ln 2, \quad (23)$$

$$-L_1(p) \leq \frac{dB}{dp}(p) \leq -U_0(p). \quad (24)$$

Further on we show that these bounds can be interpreted as benchmarks for more complex processes such as Markov processes.

3.2. Information Source of Markov Type

Consider as an Information Source the two-state Markov process. Under the notation from Section 2 we introduce the "jumping" parameter $0 \leq s \leq 2$, which in fact can be interpreted as the tendency of transition from one state to the other state

$$s := p_{0|1} + p_{1|0} . \quad (25)$$

As we show below this parameter plays a crucial role in qualitative behavior of the *IFQ* coefficient. Observe that for Markov case

$$p = \frac{p_{1|0}}{p_{0|1} + p_{1|0}} = \frac{p_{1|0}}{s} \quad (26)$$

and for $0 \leq s \leq 1$ the firing frequency p is in the full interval $[0, 1]$, while for $1 < s < 2$ it is limited to the smaller interval $1 - \frac{1}{s} < p < \frac{1}{s}$, i.e.

$$\frac{1}{2} - \left(\frac{1}{s} - \frac{1}{2}\right) \leq p \leq \frac{1}{2} + \left(\frac{1}{s} - \frac{1}{2}\right) . \quad (27)$$

thus, p is localized symmetrically around $\frac{1}{2}$. Note that for $s > 1$ the spike probability p is well separated from zero.

For the Markov source we denote the *IFQ* indicator by M . Using (11) and (12) we have

$$M := \frac{ITR}{P_{eq}(1)} = \frac{P_{eq}(0)H_1(p_{1|0}) + P_{eq}(1)H_1(p_{0|1})}{P_{eq}(1)} . \quad (28)$$

Making use of (26) we express M in terms of p and s ,

$$M_s(p) = \frac{(1-p)H_1(ps) + pH_1((1-p)s)}{p} \quad (29)$$

where H_1 is given by (3).

Observe that for $s \leq 1$ there are the following limits

$$\lim_{p \rightarrow 0} M_s = +\infty \quad \text{and} \quad \lim_{p \rightarrow 1} M_s = 0$$

and for $s > 1$

$$\lim_{p \rightarrow 1 - \frac{1}{s}} M_s(p) = \frac{[(1-s)\ln(s-1) - (2-s)\ln(2-s)]}{s-1} = \frac{H_1(s-1)}{s-1} , \quad (30)$$

$$\lim_{p \rightarrow \frac{1}{s}} M_s(p) = [(1-s) \ln(s-1) - (2-s) \ln(2-s)] = H_1(s-1) . \quad (31)$$

Next we evaluate the derivative of the quotient $M_s(p)$ over the firing rate p

$$\frac{d}{dp} M_s(p) = s \ln(sp) - s \ln[1 - (1-p)s] - s \ln(1-sp) + s \ln(s(1-p)) + \frac{\ln(1-ps)}{p^2} . \quad (32)$$

Making use of (16) and (17) we obtain the following bounds

$$4(1-p)s(2-s) \ln 2 \leq M_s(p) \leq \frac{\ln 2}{p} \quad (33)$$

and referring to the Bernoulli bounds (20) and (21) we arrive to the following limits

$$s(2-s)L_0(p) \ln 2 \leq M_s(p) \leq U_0(p) \ln 2 . \quad (34)$$

Notice that the left bound is maximal for $s = 1$, i.e. for Bernoulli case.

Now applying again (16) and (17) to (32) we get

$$\frac{1}{s} \left[\frac{(s-1)s}{p(1-p)(1-ps)} - \frac{1}{p(1-p)} \right] \leq \frac{d}{dp} M_s(p) \leq -s \left[\frac{1}{p} - \frac{(s-1)}{[1 - (1-p)s](1-sp)} \right] \quad (35)$$

and referring again to the Bernoulli bounds (21) and (22) we obtain the following limits on derivative of IFQ for Markov sources

$$-\frac{1}{s} \left[L_1(p) - \frac{(s-1)s}{p(1-p)(1-ps)} \right] \leq \frac{d}{dp} M_s(p) \leq -s \left[U_1(p) - \frac{(s-1)}{[1 - (1-p)s](1-sp)} \right] . \quad (36)$$

Observe, that for $s = 1$, inequalities (34) and (36) reduce to inequalities (23) and (24) respectively, i.e. just to Bernoulli case. Moreover, we see that for s close to 1 the bounds of $M_s(p)$ and $\frac{d}{dp} M_s(p)$ rigorously approach the Bernoulli bounds. For (34) this observation is clear while for (36) it follows from the fact that the deviations from Bernoulli bounds, $L_1(p)$ and $U_1(p)$, contain the factor $(s-1)$.

We see that for $0 \leq s \leq 1$ the derivative is negative, $\frac{d}{dp} M_s(p) < 0$, thus $M_s(p)$ is a decreasing function of p (Fig. 1) and clearly it is significantly larger for small p . For $1 < s < 2$ case the behavior of function $M_s(p)$ is qualitatively different (Fig. 2). It is non-monotonic and it has a global maximum. What means that in this case for each s the optimal firing rate exists.

The increasing s corresponds to the increasing $p_{1|0}$ so in this case the transition from the no-spike-state (state 0) to the spike-state (state 1) occurs more and more often. This means that for larger s the neuron is more firing leading to bursting phenomena.

4. Discussion and Conclusions

In this paper we address the fundamental question in neuronal coding. We analyze the possible correspondence between "temporal" and "firing rate" coding for two qualitatively different types of Information Sources. For the first type of source it is assumed that consecutive spikes are uncorrelated, thus it is governed by the Bernoulli process. In the second case we assume that there is a short time correlation (memory) between consecutive spikes, thus we model this source by the Markov process.

For the quantitative study of the relation between temporal and rate coding we propose the Information-Firing-Quotient being the ratio of Information Transmission Rate and firing rate. Since the energy used for transfer of information is proportional to the firing rate this quotient is understood as amount of information transmitted at the cost of unit energy. Clearly, for larger IFQ the transmission is more efficient. The goal is to find the optimal parameters of transmission. We found that the crucial role in qualitative and quantitative behavior of IFQ is played by the parameter s which, in fact, measures the ability of transition from non-spike to spike state and vice versa. Taking into account that in the real biological systems the firing rate is limited from below by the spontaneous activity and very small values of p and large values of IFQ , are non-realistic. This situation may happen for $s \leq 1$ (Fig. 1), when IFQ is monotonically decreasing with p , and then the realistic cut-off of p separating it from zero should be assumed. On the other hand for $s > 1$, i.e. for more active, say bursting, neurons we observe that the global maximum of IFQ exists, thus in this case there is the unique optimal firing rate (Fig. 2) well separated from zero. This leads to the non-intuitive hypothesis that even for bursting

phenomenon there may still exist the optimal regime of transmission.

Acknowledgements

We gratefully acknowledge financial support from the Polish National Science Centre under grant no. 2012/05/B/ST8/03010.

References

References

- [1] Adrian ED. (1926) The impulses produced by sensory nerve endings: Part I, *J. Physiol.*, 61, 49–72.
- [2] Ames III A, (2000) CNS energy metabolism as related to function, *Brain Res. Rev.* 34, 42–68.
- [3] Barlow HB, (1961) Possible principles underlying the transformation of sensory messages, In: Rosenblith W (ed) *Sensory Communication*. MIT Press, Cambridge.
- [4] Bialek W, Rieke F, de Ruyter van Steveninck RR, Warland D, (1991), Reading a neural code, *Science* 252, 1854–1857.
- [5] Borst A, Theunissen FE, (1999) Information theory and neural coding, *Nat. Neurosci.* 2, 947–957.
- [6] Burns BD, Webb AC, (1976) The spontaneous activity of neurons in the cat’s visual cortex, *Proc. R. Soc. Lond. B. Biol. Sci.* 194(1115), 211–23.
- [7] Butts DA, Weng C, Jin JZ, Yeh CI, Lesica NA, Alonso JM and et al., (2007) Temporal precision in the neural code and the time scales of natural vision, *Nature* 449, 92–95.
- [8] Coop AD, Reeke GN, (2001), Deciphering the neural code: neuronal discharge variability is preferentially controlled by the temporal distribution of afferent impulses, *Neurocomputing* 38–40, 153–157.

- [9] Di Maio V, (2008), Regulation of information passing by synaptic transmission: A short review, *Brain Research* 1225, 26–38.
- [10] Duguid I, Sjostrom PJ, (2006) Novel presynaptic mechanisms for coincidence detection in synaptic plasticity, *Cur. Opin. Neurobiol.* 16, 312–332.
- [11] Gerstner W, Kreiter AK, Markram H, Herz AV, (1997) Neural codes: firing rates and beyond, *Proc. Natl. Acad. Sci. U. S. A.* 94, 12740–12741.
- [12] Feller W, (1958) *An Introduction to Probability Theory and Its Applications*, United States of America: A Wiley Publications in Statistics, New York.
- [13] Goldberg DH, Andreou AG, (2004), Spike communication of dynamic stimuli: rate decoding versus temporal decoding, *Neurocomputing* 58–60, 101–107.
- [14] Hubel DH, Wiesel TN, (1962) Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex, *J. Physiol. (Lond.)* 160, 106–154.
- [15] Lansky P, Sacerdote L, (2001) The Ornstein Uhlenbeck neuronal model with the signal dependent noise, *Phys. Let. A* 285, 132–140.
- [16] Moujahid A, d’Anjou A, (2012) Metabolic efficiency with fast spiking in the squid axon, *Frontiers in Computational Neuroscience* 6, 1–8.
- [17] Paprocki B., Szczepanski J., (2013) How do the amplitude fluctuations affect the neuronal transmission efficiency, *Neurocomputing* 104, 50–56.
- [18] Pregowska A, Szczepanski J, Wajnryb E, (2015) Mutual information against correlations in binary communication channel, *BMC Neuroscience* 16(32), 1–7.
- [19] Ricciardi LM, Sacerdoteb L, (1979) The Ornstein Uhlenbeck process as a model of neuronal activity, *Biol. Cybernet.* 35, 1–9.

- [20] Richmond BJ, Optican LM Spitzer H, (1990) Two dimensional patterns are represented in temporally modulated activity of striate cortex cell, *J. Neurophysiol.* 64(2), 351–69.
- [21] Rieke F, Warland D, de Ruyter van Steveninck RR, Bialek W, (1997) *Spikes. Exploring de Neural Code*, MIT Press, Cambridge, MA.
- [22] Shannon CE, Weaver W, (1963) *The Mathematical Theory of Communication*, United States of America: University of Illinois Press, Urbana.
- [23] Stein RB, (1965) A theoretical analysis of neuronal variability, *Biophys. J.* 5, 173–195.
- [24] van Hemmen JL, Sejnowski T, (2006) *23 Problems in Systems Neurosciences*, Oxford University Press, Oxford.
- [25] Yu Q, Tang H, Tan KCh, Yu H, (2014) A brain-inspired spiking neural network model with temporal encoding and learning, *Neurocomputing* 138, 3–13.
- [26] Ash RB, (1965), *The mathematical theory of communication*, John Wiley and Sons, New York, London, Sydney, United States of America.
- [27] Cover TM, Thomas JA, (1991) *Elements of Information Theory*, A Wiley-Interscience Publication, New York, United States of America.
- [28] Arnold M, Szczepanski J, Montejo N, Wajnryb E, Sanchez-Vives MV, (2013) Information content in cortical spike trains during brain state transitions, *Journal of Sleep Research* 2, 13–21.
- [29] Strong SP, Koberle R, de Ruyter van Steveninck RR, Bialek W, (1998) Entropy and Information in Neural Spike Trains, *Physical Review Letters* 80(1), 197–200.
- [30] Lempel A, Ziv J, (1976) On the Complexity of Individual Sequences, *IEEE Transactions on Information Theory* 22, 75–81.

- [31] Kontoyiannis I, Algoet PH, Suhov YM, Wyner AJ, (1998) Nonparametric Entropy Estimation for Stationary Processes and Random Fields, with Applications to English Text, *IEEE Transactions on Information Theory* 44, 1319–1327.
- [32] Amigo JM, Keller K, Unakafova VA, (2015) Ordinal symbolic analysis and its application to biomedical recordings, *Philos. Trans. a Math. Phys. Eng. Sci.*, 373(2034), pii: 2014009.
- [33] Van Kampen NG, (2007) *Stochastic Processes in Physics and Chemistry*, North-Holland Personal Library.
- [34] Gerstner W, Kistler WM, Naud R, Paninski L, (2014) *Neuronal Dynamics, From single neurons to networks and models of cognition*, Cambridge University Press, Cambridge.
- [35] Mukherjee P, Kaplan E, (1995) Dynamics of neurons in the cat lateral geniculate nucleus: in vivo electrophysiology and computational modeling, *J. Neurophysiol.* 74(3), 1222–1243.
- [36] Reinagel P, Godwin D, Sherman SM, Koch C, (1999) Encoding of visual information by LGN bursts, *J. Neurophysiol.* 81(5), 2558–69.